duplex

Gene, 40 (1985) 169–173
Elsevier
GENE 1476

Universal restriction endonucleases: designing novel cleavage specificities by combining adapter oligodeoxynucleotide and enzyme moieties

(Recombinant DNA; predetermined cut sites; class IIS nuclease; bacteriophage M13; cloning vector; *Fok* I; hairpin sequence)

Waclaw Szybalski

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706 (U.S.A.) Tel. (608)262-1259

(Received September 3rd, 1985)
(Accepted October 2nd, 1985)

SUMMARY

Class IIS restriction endonucleases cleave double-stranded (ds) DNA at precise distances from their recognition sequences. A method is proposed which utilizes this separation between the recognition site and the cut site to allow a class IIS enzyme, e.g., *Fok* I, to cleave practically any predetermined sequence by combining the enzyme with a properly designed oligodeoxynucleotide adapter. Such an adapter is constructed from the constant recognition site domain (a hairpin containing the ds sequence, e.g., $\frac{\text{GGATG}}{\text{CCTAC}}$ for *Fok* I) and a variable, single-stranded (ss) domain complementary to the ss sequence to be cleaved (at 9 and 13 nucleotides on the paired strands from the recognition sequence in the example of *Fok* I). The ss sequence designated to be cleaved could be provided by ss phage DNA (e.g., M13), gapped ds plasmids, or supercoiled ds plasmids that were alkali denatured and rapidly neutralized. Combination of all three components, namely the class IIS enzyme, the ss DNA target sequence, and the complementing adapter, would result in target DNA cleavage at the specific predetermined site. The target ss DNA could be converted to the precisely cleaved ds DNA by DNA polymerase, utilizing the adapter oligodeoxynucleotide as primer. This novel procedure represents the first example of changing enzyme specificity by synthetic design. A practically unlimited assortment of new restriction specificities could be produced. The method should have many specific and general applications when its numerous ramifications are exploited.

INTRODUCTION

There exist more than 500 class II restriction endonucleases that cleave DNA at specific sites. The number of different recognition sites is greater than 100 (not counting various modifications, including methylation) (Roberts, 1985; Kessler et al., 1985). Although the existing repertoire of cleavage sites is of great value, there is frequently a need for additional cut sites to permit precise engineering of genes and

Abbreviations: bp, base pair(s); class IIS enzymes; class II restriction endonucleases that cut at precise sites 2–13 nt away from their 4–7-bp recognition site (see Table I); ds, double stranded; nt, nucleotide(s); Pollk, Klenow fragment of *E. coli* DNA polymerase I; shifters, class IIS enzymes; ss, single stranded; UV, ultraviolet light.

regulatory sequences. At present, to access a sequence where a favorable restriction site is not available, one usually engineers the sequence by a variety of methods, including insertion of appropriate linkers, mutagenesis, or exonucleolytic digestion. However, such methods for creating a new site modify the nucleotide sequence under study and may often affect its function.

I propose here an alternative general approach which leaves the sequence unaltered while modifying the specificity of a restriction enzyme to precisely cleave at any predetermined site. An example is given.

PRINCIPLE OF THE METHOD

More than ten restriction enzymes of class II, for which the cleavage site is shifted a considerable (up to 13 nt) but precise distance from the recognition site, have been described. I shall refer to such endonucleases as 'shifters' or 'class IIS' enzymes. Their recognition sites and the displacements of their cleavage sites are listed in Table I.

I propose to engineer these class IIS restriction endonucleases by using synthetic oligodeoxynucleotide adapter molecules to instruct the enzyme to recognize any predetermined sequence. Fig. 1

TABLE I

Properties of class IIS restriction endonucleases ('shifters')

Enzyme	Recognition sequences [a]		Protruding end [b]	Reference [c]	
				Page	Entry
*Bbv*I[d,e]	5′–GCAGCN$_8$↓ CGTCGN$_{12}$↓	5′–↓N$_{12}$GCTGC ↓N$_8$ CGACG	5′N$_4$	57	513
*Bbv*II	5′–GAAGACN$_2$ CTTCTGN$_6$	5′–N$_6$GTCTTC N$_2$CAGAAG	5′N$_4$	102	–
*Bin*I	5′–GGATCN$_4$ CCTAGN$_5$	5′–N$_5$GATCC N$_4$CTAGG	5′N$_1$	57	516
*Fok*I[d,e]	5′–GGATGN$_9$ CCTACN$_{13}$	5′–N$_{13}$CATCC N$_9$ GTAGG	5′N$_4$	57	519
*Hga*I[d,f]	5′–GACGCN$_5$ CTGCGN$_{10}$	5′–N$_{10}$GCGTC N$_5$ CGCAG	5′N$_5$	58	526
*Hph*I[d,e]	5′–GGTGAN$_8$ CCACTN$_7$	5′–N$_7$TCAGG N$_8$AGTCC	3′N$_1$	58	527
*Mbo*II[d,e]	5′–GAAGAN$_8$ CTTCTN$_7$	5′–N$_7$TCTTC N$_8$AGAAG	3′N$_1$	58	528
*Mnl*I[d,f]	5′–CCTCN$_7$ GGAGN$_7$	5′–N$_7$GAGG N$_7$CTCC	blunt	59	531
*Sfa*NI[d,f]	5′–GCATCN$_5$ CGTAGN$_9$ G G	5′–N$_9$GATGC N$_5$CTACG G G	5′N$_4$	59	532
*Taq*II	5′–CACCCAN$_{11}$ CTGGGTN$_9$ G G	5′–N$_9$ TCGGTC N$_{11}$ACCCAG G G	3′N$_2$	59	533
*Tth*111II[g]	5′–CAAPuCAN$_{11}$ GTTPyGTN$_9$	5′–N$_9$ TGPyTTG N$_{11}$ACPuAAC	3′N$_2$	59	535

[a] Since the recognition sites are not palindromic, both orientations are listed. Arrows specify the cleavage sites (shown only for the first enzyme).

[b] The 5′ or 3′ protruding ss ends of the restriction sites are indicated and their lengths given.

[c] Refers to the review of Kessler et al. (1985), with the page and entry number specified.

[d] Commercially available (e.g., New England Biolabs; see Kessler et al., 1985).

[e] No endonucleolytic activity toward M13 ss DNA (Podhajska and Szybalski, 1985).

[f] Displays endonucleolytic activity toward M13 ss DNA (Podhajska and Szybalski, 1985).

[g] Py, T or C; Pu, A or G. For cut points see Roberts (1985).

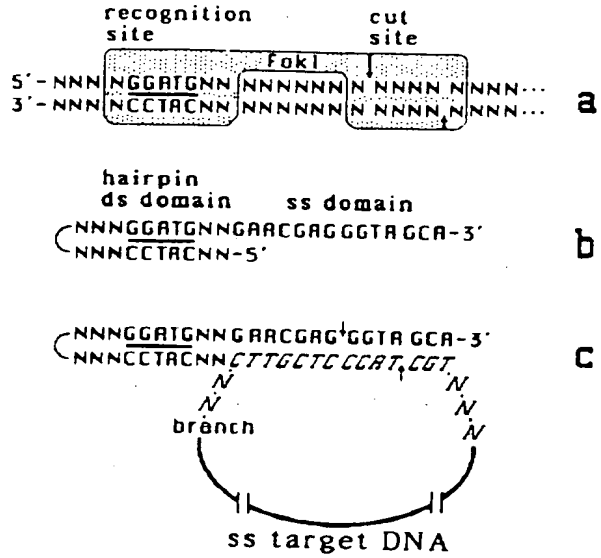

Fig. 1. Design of an adapter oligodeoxynucleotide for instructing the *Fok*I enzyme to cleave M13 ss DNA at a predetermined point. (a) The *Fok*I recognition site, $\frac{\text{GGATG}}{\text{CCTAC}}$ and the cut site for ds DNA. Both the recognition and cutting domains of the enzyme are schematically outlined (stippled). The staggered cleavage points are represented by arrows. (b) The 34-mer adapter oligodeoxynucleotide designed to react with *Fok*I and to cut M13 ss DNA between nt 1341 and 1342 (see Van Wezenbeek et al., 1980). Symbol N represents any deoxynucleotide. (c) Base pairing between the adapter and M13 ss DNA. M13 ss DNA (ss target DNA) is represented by italicized letters connected by the broken curved line. The only steric differences between the (a) and (c) sequences in contact with the *Fok*I enzyme are (i) the branch (whisker) starting at *C* (nt 1352 of M13 DNA) and (ii) the ss discontinuity between ...NN and the above *C*. *Fok*I activity should not be affected by this steric modification, since it is outside the *Fok*I recognition and cut sites and within a region of nonspecific sequence requirement. Only one kind of adapter with the 5' hairpin ds domain (constant domain) and with the 3' ss target-recognition domain (variable domain) is shown here. The second kind of adapter with the 3' hairpin ds domain and 5' ss variable domain is not presented here, but is shown in Fig. 1a of the accompanying report of Podhajska and Szybalski (1985).

outlines the principle of the method as applied to the *Fok*I enzyme. The relationship between the recognition site, $\frac{\text{GGATG}}{\text{CCTAC}}$ and the staggered cleavage sites on the normal ds DNA molecule is shown in Fig. 1a. The adapter oligonucleotide (Fig. 1b) that mimics the above arrangement consists of the hairpin (ds) domain, which carries the *Fok*I recognition sequence (underlined), and an ss domain where the cleavage(s) would occur after pairing with the complementary target DNA strand. In Fig. 1c, a segment of M13 phage ss DNA represents the target DNA strand. Thus, the ds domain of the adapter is a constant part that provides specificity for the IIS enzyme, whereas the ss domain is variable and could be designed to provide cleavage at any chosen site in the complementary target DNA.

In Fig. 1, only one kind of adapter, with the 3' ss domain, is shown. Another adapter, with the 5' ss domain is not presented here, but both kinds of adapters were synthesized and are discussed in the accompanying communication (Podhajska and Szybalski, 1985).

PROPOSED PROCEDURES

Although the experimental procedures are in the developmental stage (see Podhajska and Szybalski, 1985), various aspects of the technique are outlined here.

(a) Design of the adapter

The oligodeoxynucleotide depicted in Fig. 1b will serve as an example of adapter design for the *Fok*I enzyme and ss target DNA from the M13 phage. The hairpin ds domain of this 34-mer contains the $\frac{5'\text{–GGATG}}{3'\text{–CCTAC}}$ recognition site for *Fok*I; its ss domain was designed to direct the cleavage between T and C in the target M13 DNA. The complex between the adapter and M13 DNA (Fig. 1c) mimics a ds module (Fig. 1a) which contains both the *Fok*I recognition site and the ds cleavage site (see Table I and Fig. 1c). Depending on the sequence of the ss domain of the adapter, the enzyme would cut at any chosen site on the target DNA. At the same time, the 3' end of the adapter (see Fig. 1b,c) could serve as primer for synthesis of the strand complementary to the target ss DNA, possibly in two stages: a limited synthesis with three dNTP's (to anchor the 3' end of the adapter before *Fok*I digestion, but not so extensive as to provide ds recognition sites for *Fok*I), followed by synthesis of the total complementary strand after the *Fok*I cleavage.

Fig. 1 does not indicate optimal adapter design. Optimization would include: (i) determination of the

most efficient hairpin lengths to the left and to the right of the GGATG recognition site; (ii) alteration of the length of the ss domain complementary to the cleaved ss DNA; (iii) modifying the position of the ss M13 DNA branch (Fig. 1b) with respect to the GGATG and cleavage sites; (iv) comparison of adapters with the 5′ and 3′ ss domains (see Fig. 1 in the accompanying report of Podhajska and Szybalski, 1985); (v) possible use of ambiguous nucleotides (inosinic acid) or ribonucleotides in the ss domain of the adapter; and (vi) inclusion of reactive deoxynucleotides (e.g., 5-iododeoxyuridine nucleotides) in the tip of the hairpin in order to form covalent (but active) adapter-enzyme complexes upon UV irradiation. Optimization of adapter design should also lead to a better understanding of the mechanism of the endonucleolytic activity of the class IIS enzymes.

Although it would be necessary to design a specific adapter for each chosen cleavage site, only the variable ss domain (5-15 nt) would have to be custom-made. It could then be either (i) ligated to the stem of the constant ds hairpin domain, using as a splicing support a set of oligodeoxynucleotides designed to be complementary to the adjoining sequences of the hairpin and ss domains or, (ii) prepared by first synthesizing a supply of the constant hairpin domain with protective groups retained, and in the second synthetic stage adding the variable ss domain. Such two-step organic synthesis is possible for adapters containing the 5′ ss domain (see Fig. 1a in the accompanying report of Podhajska and Szybalski, 1985), because of the 3′ → 5′ direction of the nucleotide coupling.

Although the adapter synthesis would have to be customized for many sequences, one could develop an inventory of adapters for large numbers of commonly encountered sequences for which no enzymes are presently available.

(b) Target DNA and vectors

(1) Single-stranded DNA

One of the main advantages of using ss target-vector DNA, as exemplified by M13 ss DNA shown in Fig. 1., is that it does not provide recognition sites for most of the IIS enzymes, which are rather frequent cutters of ds vectors. Fortunately, several commercially available class IIS enzymes (*Fok* I, *Hph* I, *Mbo* I, *Bbv* I; New England Biolabs) do not have interfering endonucleolytic activity toward M13 DNA (Podhajska and Szybalski, 1985). The ss target DNA could be obtained from a variety of sources, including the ss DNA phage vectors (M13, fd, f1) or any ds plasmid carrying the replication origin (*ori*) of those phages that convert ds plasmids to packaged ss DNA upon phage superinfection.

(2) Double-stranded DNA with ss gaps

Another substrate could be ds DNA with properly designed ss gaps, but in this case the ds cleavage sites should preferably be either absent or rendered inactive by (i) methylation or other modification of the ds DNA, or (ii) saturation of enzyme with the adapter and possible covalent binding (see section a above). For the specific methylation of *Fok* I sites, a *Fok* I methylase is available (New England Biolabs). For class IIS enzymes which create 4-nt cohesive ends, the religation of the cleaved plasmid would restore its original arrangement, since each cleavage sequence is unique. However, a cut in the ss gap would remain open, especially if adapter molecules were removed prior to religation.

(3) Double-stranded DNA

In addition to the problems specified in section b2 above, use of a ds DNA substrate requires a special method for threading the ss domain of the adapter into the ds target sequence. This should be possible when using reversibly denatured target DNA obtained by alkali denaturation of the supercoiled plasmid DNA followed by rapid neutralization. Such a procedure results in largely denatured molecules shown to anneal with primer oligonucleotides (Chen and Seeburg, 1985). Use of a single adapter would result in an ss nick, whereas two adapters with barely overlapping or neighboring ss domains complementary to opposing strands should provide staggered cuts. Alternatively, this might be accomplished by use of catalysts, e.g., RecA-like protein, mild denaturation conditions, or of ribonucleotides in the ss domain of the adapter. There are no *a priori* reasons why this could not be accomplished with some ingenious design. Also, with the adapter shown in Fig. 1, the ds DNA could be obtained by using the 3′-OH end of the adapter as a primer to synthesize the complementary strand, as discussed in section a above. Employing this kind

of approach, we have been able to prepare ds DNA fragments with tailored end(s), ready for cloning (A.J. Podhajska and W.S., in preparation).

DISCUSSION

(a) Design of novel enzyme specificities

The proposed procedure is intended to permit construction of novel cleavage specificities by combining carefully contrived synthetic adapter molecules with class IIS restriction endonucleases. This appears to be the first example of the creation of new enzyme specificities designed at will. Devising unlimited numbers of new enzymatic specificities ('new restriction enzymes') should be limited solely by oligodeoxynucleotide synthesizer capacity.

(b) Use of class IIS enzymes with tailored specificities

This proposal pertains mainly to ss DNA and describes how to cut it at any specified site using the adapter molecule. However, extending the 3'-OH end of the adapter (see Fig. 1c) with PolIk (see PROPOSED PROCEDURES, section **a**) would create a ds molecule with precise end(s). For *Fok* I, a pairing between approx. 9 nt, including both cuts, is probably all that is necessary for a successful cleavage (see Fig. 2, cut II, in the report of Podhajska and Szybalski, 1985). Thus, any particular adapter would seek specific sites and create identical but asymmetric 4-nt cohesive ends, permitting the design of specific constructs. Whereas native *Fok* I enzyme creates a variety of 4-nt cohesive ends, each *Fok* I-adapter complex would produce a unique 4-nt cohesive sequence.

(c) Tailoring the specificity of other enzymes

Although only the class IIS restriction endonucleases are considered in this proposal, it should be possible to design adapters for enzymes that recognize an inverted repeat separated by a run of nonspecific base pairs. For instance, the adapter $\subset$NCCANNNNNNTGGN$\supset$ / $\subset$NGGT ACCN$\supset$ should convert the *Bst*XI enzyme (see Kessler et al., 1985) to an endonuclease seeking an ss sequence complementary to the central NNNNNN ss domain of the adapter. Many other kinds of adapters could be designed.

(d) Conclusions

The purpose of this proposal is to provide a novel method for designing new restriction enzyme specificities by the use of purposefully designed adapter molecules. A model for the application of this new principle and a few ramifications are outlined here. An example of the practical application of this method is described in the accompanying communication (Podhajska and Szybalski, 1985). Detailed procedures as well as other applications and ramifications of the principle will be presented elsewhere.

ACKNOWLEDGEMENTS

I am thankful to Elizabeth H. Szybalski, Bette Sheehan, Ilse Riegel, Van Potter, A. Jagoda Podhajska and Mick Noordewier for critical reading and constructive comments.

REFERENCES

Chen, E.Y. and Seeburg, P.H.: Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA 4 (1985) 165-170.

Kessler, C., Neumaier, P.S. and Wolf, W.: Recognition sequences of restriction endonucleases and methylases – a review. Gene 33 (1985) 1-102.

Podhajska, A.J. and Szybalski, W.: Conversion of the *Fok* I endonuclease to a universal restriction enzyme: cleavage of the phage M13mp7 DNA at predetermined sites. Gene 40 (1985) 175-182.

Roberts, R.J.: Restriction and modification enzymes and their recognition sequences. Nucl. Acids Res. 13 (1985) r165-r200.

Van Wezenbeek, P.M.G.F., Hulsebos, T.J.M. and Schoenmakers, J.G.G.: Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd. Gene 11 (1980) 129-148.

Communicated by Z. Hradečná.